

CLAIMS

1 1. A method for operating a server cluster where service requests to a first server are taken
2 over by a second server when a system operator initiates the takeover from the first server, and
3 both servers have storage means, the method comprising the steps of:

4 sending a first indication from the first server to the second server requesting the second
5 server to take over the operations of the first server;

6 sending a second indication from the second server to the first server requesting the first
7 server to shutdown; and

8 shutting down the first server by completing service requests that were being processed
9 when the second indication was received at said first server, and before said second server takes
10 over the storage means and operations of said first server.

1 2. The method in accordance with claim 1 further comprising the step of sending a periodic
2 request from the second server to the first server to stay shut down, after the first server has shut
3 down.

1 3. The method in accordance with claim 1 wherein said servers issues periodic signals to
2 each other indicating they are functioning at some level, and said second server does not issue
3 said second indication to said first server when the periodic signals issued by said first server are
4 absent, and said second server instead issues a signal to immediately takeover the storage means
5 and operations of said first server..

1 4. A method for operating a server cluster where service requests to a first server are taken
2 over by a second server when a system operator initiates the takeover from the second server,
3 and both servers have storage means, the method comprising the steps of:

4 sending a second indication from the second server to the first server requesting the first
5 server to shutdown; and

6 shutting down the first server by completing service requests that were being processed when the
7 second indication was received at said first server, and before said second server takes over the
8 storage means and operations of said first server

1 5. The method in accordance with claim 1 further comprising the step of monitoring the
2 status of the first server to detect any operational faults prior to sending of the first indication to
3 the second server in the event an operational fault is detected.

1 6. The method in accordance with claim 1 further comprising the step of letting the first
2 server complete certain functions it is performing at the time the operational fault is detected
3 before it shuts down.

1 7. The method in accordance with claim 6 further comprising the step of sending a periodic
2 request from the second server to the first server to stay shut down, after the first server has shut
3 down.

1 8. A method for operating a first server to provide backup service to a second server when
2 the first server detects an operational fault, the method comprising the steps of:

3 receiving a first request at the second server requesting the second server to take over the
4 operations of the first server when the first server has detected an operational fault that will
5 require it to shut down;

6 permitting the first server to complete certain functions it is performing at the time the
7 operational fault is detected before it shuts down; and

8 taking over the functions of the first server at the second server after the first server has
9 shut down.

1 9. The method in accordance with claim 8 further comprising the step of monitoring the
2 operational status of the first server to detect any operational faults and cause sending of the first
3 request to the second server to take over the operations of the first server in the event an
4 operational fault is detected on the first server.

1 10. The method in accordance with claim 9 further comprising the steps of:
2 determining if the second server can provide backup service to the first server; and
3 requesting the first server to shut down if the second server can provide backup service to
4 the first server.

1 11. The method in accordance with claim 10 further comprising the step of receiving an
2 indication at the second server from the first server indicating the type of operational fault
3 detected by the first server.

1 12. The method in accordance with claim 11 further comprising the step of receiving a
2 periodic request at the first the first server sent from the second server for the first server to stay
3 shut down.

1 13. The method in accordance with claim 12 further comprising the step of rebooting the first
2 server after the detected operational fault has been cured.

1 14. Apparatus for transferring service requests from a first server to a second server in a
2 cluster of servers when the first server has a fault therein, said apparatus comprising:

3 first means in said first server for monitoring the operation of said first server, said first
4 monitoring means generating a first fault signal when it is determined that said first server has a
5 fault therein, said first fault signal being sent to said second server; and

6 wherein when said second server receives said first fault signal it permits said first server
7 to complete existing service request that are being processed by said first server and then takes
8 over and processes further service requests directed to said first server.

1 15. Apparatus for transferring service requests from a first server to a second server in a
2 cluster of servers when the first server has a fault therein, and said first server sends a fault signal
3 to said second server upon there being a fault in said first server, said apparatus comprising:

4 first means in said second server for receiving said fault signal from said first server; and

5 wherein when said second server receives said first fault signal it permits said first server
6 to complete existing service request that are being processed by said first server and then takes
7 over and processes further service requests directed to said first server.

1 16. The apparatus in accordance with claim 15 further comprising memory means in said
2 second server, said first server generates a compilation of the status of service requests being
3 processed by said first server and sends said compilation to said second server to be stored in
4 said memory means thereat, said compilation being used by said second server to take over
5 processing of service requests directed to said first server in an orderly fashion.

1 17. The apparatus in accordance with claim 16 wherein after said second server takes over
2 processing of service requests directed to said first server, said second server transmits stay dead
3 signals to said first server causing it not to attempt to commence processing new service
4 requests.

1 18. The apparatus in accordance with claim 17 wherein said first and said second servers are
2 file servers.

1 19. A computer readable medium containing executable instructions for providing backup
2 service between a first server and a second server operating in a cluster mode when the first
3 servers experiences a fault that will cause it to shut down, the executable program instructions
4 comprising program instructions for:

5 sending a fault signal from the first server to the second server;

6 determining if the second server can provide backup service to the first server;

7 requesting the first server to shut down if the second server can provide backup service to
8 the first server; and

9 taking over the functions of the first server by the second server after the first server has
10 complete current operations and has shut down.

- 1 20. Apparatus for operating a first server to receive backup service from a second server
2 when the first server detects an operational fault, said apparatus comprising:
3 means for sending a first indication from the first server to the second server when the
4 first server detects the operational fault that will require it to shut down;
5 means for sending a second indication from the first server to the second server
6 indicating the type of operational fault detected by the first server;
7 means for receiving a shutdown command from the second server at the first server if the
8 second server can provide backup service to the first server; and
9 means for shutting down the first server.
- 1 21. Apparatus for operating a first server to provide backup service to a second server when
2 the first server detects an operational fault, said apparatus comprising:
3 means for receiving a first request at the second server from the first server requesting the
4 second server to take over the operations of the first server when the first server has detected an
5 operational fault that will require it to shut down;
6 means for permitting the first server to complete certain functions it is performing at the
7 time the operational fault is detected before it shuts down; and
8 means for taking over the functions of the first server at the second server after the first
9 server has shut down.
